

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-194612

(43)Date of publication of application : 14.07.2000

(51)Int.Cl.

G06F 13/00

G06F 12/00

H04Q 7/38

H04L 29/06

(21)Application number : 11-364432

(71)Applicant : SPYGLASS INC

(22)Date of filing : 22.12.1999

(72)Inventor : CARLINO KENNETH F
HOHL KENNETH B

(30)Priority

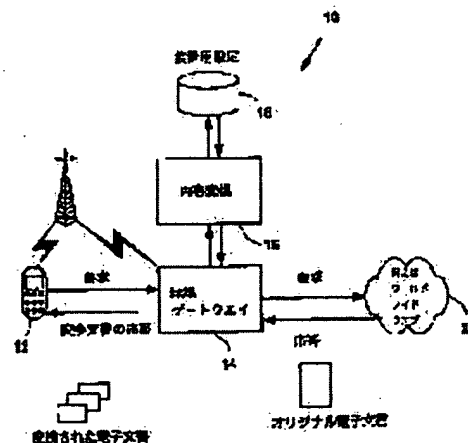
Priority number : 98 221750 Priority date : 28.12.1998 Priority country : US

(54) METHOD AND DEVICE FOR CONVERTING CONTENTS OF ELECTRONIC DATA INTO CONTENTS USED FOR RADIO DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a method and system to convert the contents of electronic data into those which are used for a wireless device, i.e., to enable the wireless device such as a wireless telephone set to receive an electronic document including the worldwide Web pages, etc., of an internet and the electronic data in a format that is most suitable for displaying the document.

SOLUTION: An original electronic document which is written in a 1st markup language such as an HTML (hypertext markup language) is converted into a 2nd WML(wireless markup language), etc., of a WAP(wireless application protocol). The text document elements and non-text document elements (e.g. images) of the original electronic document are converted into a format suitable to a 2nd markup language from a format suitable to the 1st markup language. Then a converted document (e.g. WML) suitable to the display on a wireless device is transmitted as a reply to the request given to the original document (e.g. HTML). Thus, a user is more satisfied with the wireless device.



LEGAL STATUS

[Date of request for examination] 22.12.1999

[Date of sending the examiner's decision of rejection] 03.06.2003

[Kind of final disposal of application other than the examiner's decision of rejection or

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-194612

(P2000-194612A)

(43) 公開日 平成12年7月14日 (2000.7.14)

(51) Int.Cl. ⁷	識別記号	F I	テマコード [*] (参考)
G 0 6 F 13/00	3 5 1	G 0 6 F 13/00	3 5 1 B
	5 4 6	12/00	5 4 6 R
H 0 4 Q 7/38		H 0 4 B 7/26	1 0 9 M
H 0 4 L 29/06		H 0 4 L 13/00	3 0 5 B

審査請求 有 請求項の数34 O L (全 31 頁)

(21) 出願番号 特願平11-364432

(22) 出願日 平成11年12月22日 (1999. 12. 22)

(31) 優先権主張番号 0 9 / 2 2 1 7 5 0

(32) 優先日 平成10年12月28日 (1998. 12. 28)

(33) 優先権主張国 米国 (US)

(71) 出願人 599109939

スパイグラス インコーポレイテッド

Spyglass, Inc.

アメリカ合衆国 60563 イリノイ州 ネ

イバーヴィル イー. ディール ロード

1240

(72) 発明者 ケネス エフ. カルリーノ

アメリカ合衆国 60188 イリノイ州 キ

ャロル ストリーム ナラガンセット

1228

(74) 代理人 100083806

弁理士 三好 秀和 (外1名)

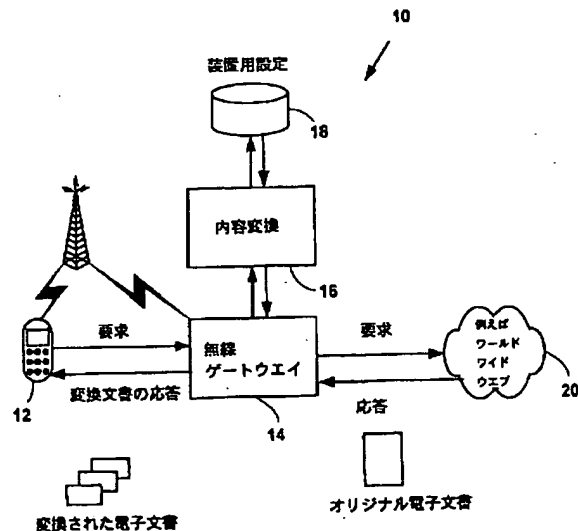
最終頁に続く

(54) 【発明の名称】 電子データの内容を無線装置用に変換する方法およびシステム

(57) 【要約】

【課題】 本発明は、電子データの内容を無線装置用に変換するための方法およびシステムを提供することを課題とする。すなわちインターネットのワールドワイドウェブのウェブページ等、電子データ含む電子文書を、無線電話等の無線装置がその表示に適したフォーマットで受け取れるようにすることを課題とする。

【解決手段】 本発明は、ハイパーテキストマークアップ言語 (HTML) 等の第1のマークアップ言語で書かれたオリジナル電子文書を、無線アプリケーションプロトコル (WAP) の無線マークアップ言語 (WML) 等の第2の無線マークアップ言語に変換する。オリジナル電子文書のテキスト文書要素および非テキスト文書要素 (例えば画像) は、第1のマークアップ言語に適したフォーマットから第2のマークアップ言語に適したフォーマットに変換する。オリジナル電子文書 (例えばHTML) に対する要求の応答として、無線装置での表示に適するように変換した文書 (例えばWML) を送出することにより、無線装置のユーザの満足度を向上させる。



【請求項16】 前記変換後の電子文書を送る段階が、前記第1の無線装置が変換後の電子文書の全部を格納するのに十分なメモリを有するか否かを決定する段階と、前記第1の無線装置が十分なメモリを持っていないければ、前記変換後の電子文書を複数のサブ文書に分割し、前記複数のサブ文書を前記変換後の電子文書のすべてを送る代わりに送る段階とを備えることを特徴とする請求項1に記載の方法。

【請求項17】 複数のネットワーク装置を有するネットワークにおいて電子文書を分割する方法であって、
10 オリジナル電子文書を第2の無線マークアップ言語の論理的サブ文書に分割するにあたり、テキスト要素と非テキスト要素とを含む第1のセットの文書要素を第1のマークアップ言語において選択する段階と、
前記第1のセットの文書要素に基づき、前記オリジナル電子文書を文書要素に分解する段階と、
前記分解したオリジナル電子文書の文書要素に基づき、
前記オリジナル電子文書を、前記分解したオリジナル電子文書の文書要素の論理的順序を保持させながら、第2
10 の無線マークアップ言語の複数のサブ文書に分割する段階とを備えることを特徴とする電子文書を分割する方法。

【請求項18】 請求項17に記載の方法を中央演算処理装置に実行させるための命令群を格納したことを特徴とするコンピュータで読み取り可能な媒体。

【請求項19】 前記第1のマークアップ言語がHTML(ハイパーテキストマークアップ言語)であり、前記第2の無線マークアップ言語がWML(無線マークアップ言語)であることを特徴とする請求項17に記載の方法。

【請求項20】 前記第2の無線マークアップ言語の複数のサブ文書が、複数の無線マークアップ言語カードであることを特徴とする請求項17に記載の方法。
30

【請求項21】 前記第1のマークアップ言語の第1のセットの要素が、ハイパーテキストマークアップ言語における、USEMAP属性を有するIMG要素およびOBJECT要素と、NOFRAMES要素と、FRAME要素と、FIELDSET要素と、IFRAME要素と、NOSCRIPT要素と、TBODY要素と、TABLE要素とのいずれかを含むことを特徴とする請求項17に記載の方法。

【請求項22】 複数のネットワーク装置を有するネットワークにおいて電子文書を分割する方法であって、
40 (a)オリジナル電子文書を第2の無線マークアップ言語の複数のサブ文書に分割するにあたり、各カテゴリが第1のマークアップ言語の1つ以上の文書要素に対応するようにしながら第1のマークアップ言語の第1のセットのカテゴリを選択する段階と、
(b)前記オリジナル電子文書の構造化モデルを作成する段階と、
(c)前記第1のセットのカテゴリに対応して前記構造化モデルを1つ以上の文
50

書要素に分解する段階と、

(d)第2の無線マークアップ言語の第1のサブ文書に関連する第1のリストに、1つのカテゴリを追加する段階と、

(e)前記分解した文書要素の1つ以上を、前記1つのカテゴリに関連する第2の無線マークアップ言語の次のサブ文書上の次のリストに追加する段階と、

(f)前記第1のサブ文書と前記次のサブ文書との間に状態情報を含むリンクを作成する段階と、

(g)前記段階(c)から(f)までを前記オリジナル電子文書の構造化モデルの最後まで繰り返すことにより、前記オリジナル電子文書を第2の無線マークアップ言語の複数のサブ文書に分割する段階とを備えることを特徴とする電子文書を分割する方法。

【請求項23】 請求項22に記載の方法を中央演算処理装置に実行させるための命令群を格納したことを特徴とするコンピュータで読み取り可能な媒体。

【請求項24】 前記第1のマークアップ言語の第1のセットのカテゴリが、リンク、テキスト、表入力フィールド、ページ選択、画像リンク、画像、ナビゲーションのいずれかを含むことを特徴とする請求項22に記載の方法。

【請求項25】 前記第1のマークアップ言語がHTML(ハイパーテキストマークアップ言語)であり、前記第2の無線マークアップ言語がWML(無線マークアップ言語)であることを特徴とする請求項22に記載の方法。

【請求項26】 前記段階(b)が、前記オリジナル電子文書の文書オブジェクトモデルを作成する段階を含むことを特徴とする請求項22に記載の方法。

【請求項27】 前記段階(d)が、無線マークアップ言語の第1のカードに関連する第1のリストに1つのカテゴリを追加する段階を含むことを特徴とする請求項22に記載の方法。

【請求項28】 前記段階(e)が、分解した文書要素の1つ以上を、前記1つのカテゴリに関連する無線マークアップ言語の次のカード上の次のリストに追加する段階を含むことを特徴とする請求項22に記載の方法。

【請求項29】 前記段階(f)が、前記第1のサブ文書と前記次のサブ文書とを関連付けることによって、無線マークアップ言語カードデッキにおける第1のカードと次のカードとの間に無線マークアップ言語リンクを作る段階を含むことを特徴とする請求項22に記載の方法。

【請求項30】 複数のネットワーク装置を有するネットワークにおいて電子文書を分割する方法であって、オリジナル電子文書から1つ以上の文書要素を抽出するための抽出式のセットを作成する段階と、
前記抽出式のセットを前記オリジナル電子文書に適用することによってテキスト文書要素および非テキスト文書要素を含む1つ以上の文書要素を抽出する段階と、
50 前記抽出した文書要素を用いて前記オリジナル電子文書

ダウンロードしたりアップロードできる。ゴーファは、ファイル転送プロトコルと同様であり、ハイパーテキストを含んだファイルにリンクした一連のメニューを提供する。

【0009】現在の無線電話等の無線装置は、通話機能に加えデータ機能を持っている。このデータ機能により、無線装置はワールドワイドウェブから電子文書を受け取ることができる。電子文書を無線装置で使用可能なフォーマットに最適化するため、無線アプリケーションプロトコル(WAP)が一般に使われている。このプロトコルは、いくつかのプロトコルと規格を含み、無線装置に電子文書へのアクセスを提供すべく設計されている。このプロトコルは、ワールドワイドウェブ用に開発された他のマークアップ言語やプロトコルの代替として開発された。無線アプリケーションプロトコルのさらに詳しい情報は、ワールドワイドウェブの「<http://www.wapforum.org>」にある。

【0010】無線アプリケーションプロトコルの一つの要素は、無線マークアップ言語(WML)である。この無線マークアップ言語は、マークアップタグを含み、電子文書のフォーマットとレイアウトについての制御を提供する。無線マークアップ言語は、ハイパーテキストマークアップ言語等と比べ、無線電話等の無線装置での使用に適している場合が多い。

【0011】無線マークアップ言語データは、「カード」の集まりとして構成される。カードの集まりは「デッキ」と呼ばれる。各カードは、構造化された内容とナビゲーション仕様とを含む。論理的には、無線装置のユーザは、一連のカードを調べ、各カードの内容を閲覧し、要求情報を入力し、オプションを選択し、デッキ内の他のカードを調べる。

【0012】

【発明が解決しようとする課題】無線装置のユーザは、ワールドワイドウェブの電子文書を要求することもできる。しかしながら、ワールドワイドウェブの電子文書の内容を無線装置の小さなディスプレイに表示させるにはいくつかの問題がある。

【0013】まず第一に、ワールドワイドウェブ用に開発された電子文書は、標準的な「スーパーVGA」解像度(例えば800×600画素、256色以上のカラー表示)で表示されることを前提としている。無線電話等の無線装置のユーザも、ワールドワイドウェブの電子文書を見る時がある。しかしながら、ほとんどの無線装置は、スーパーVGAよりも低い解像度(例えば100×200画素、モノクロ表示)を有し、カラーをサポートしていない。また、スーパーVGAを有する装置は、80文字以上のテキストを24行以上表示できる能力を持つ場合が多い。しかるに無線装置のディスプレイは、20文字のテキストを約2～5行しか表示できないことが多い。このため電子文書の内容を適切に表示できず、ユーザは

元のハイパーテキストの内容を見ることが難しい。しかしながら電子文書は、スーパーVGA解像度のディスプレイを持つ装置にも、それ以下の解像度のディスプレイを持つ無線装置にも、同じ内容を提供する必要がある。

【0014】1つの解決策は、電子文書のいくつかのバージョンをコンテンツプロバイダのサイトに保存しておくことである。例えば電子文書の1つのバージョンはハイパーテキストマークアップ言語で保存し、他のバージョンは無線マークアップ言語で保存することである。しかしながら、この方法は大きな記憶空間を必要とする上、電子文書が変更された場合に各バージョンの同期を取るための保守作業が極めて難しい。

【0015】別の問題は、無線装置のディスプレイは、電子文書を含むテキスト、画像、動画、ビデオ等の全ページをダウンロードし表示することができない、あるいはそのための十分なメモリを持たないことである。多量のテキスト、画像、動画、ビデオ等を含む電子文書を表示しようとする、無線装置は過負荷になり、悪影響を受け、ユーザをいらさらせる。

【0016】さらに別の問題は、無線装置のほとんどのユーザは、その装置をワールドワイドウェブを幅広く閲覧するために使うのではないことである。無線装置のユーザは、例えば電話番号、住所、株価、スポーツの点数結果、最新のニュース等、特定の情報を探すことが多い。このようなユーザは、ハイパーテキスト文書のある種の電子的内容(例えば画像)を表示させないことを望む。

【0017】このため、オリジナル電子文書を無線電話等の無線装置で利用できるように変換する方法およびシステムを提供する必要がある。これら方法およびシステムは、ワールドワイドウェブのほぼすべての電子文書を無線装置に適したフォーマットの文書に変換し表示することを可能にする。

【0018】

【課題を解決するための手段】本発明は、無線装置に電子文書を表示する上での問題点を解決するため、電子文書の内容を変換する方法およびシステムを提供する。本発明の一態様は、電子文書の内容を変換する方法に関する。この方法は、第1のネットワーク上の第2のネットワーク装置にあるオリジナル電子文書に対する要求を、前記第1のネットワーク上の第1のネットワーク装置から受け取る段階を含む。前記第1のネットワーク装置は無線装置であり、前記要求は前記無線装置のタイプを示すデータを含む。この方法は、第1のマークアップ言語で書かれた前記オリジナル電子文書を、複数の電子文書要素に分解する。これら文書要素は、テキスト要素と非テキスト要素とを含み、第2の無線マークアップ言語に変換される。すなわち、第1のマークアップ言語で書かれたオリジナル電子文書のテキスト要素は、第2の無線マークアップ言語のテキスト要素に変換される。第1の

(図1に示さず)。本発明の好適実施例は、内容変換システム10に示したネットワーク要素に限定されず、より多くのあるいはより少ないネットワーク要素を使っても良い。本発明の他の実施例において、第1の無線装置12は、有線ネットワーク装置(例えばワイヤやケーブルを用いて臨時にまたは永久的にネットワークに接続したネットワーク装置)でも良い。このような有線ネットワーク装置は、1つ以上の無線マークアップ言語(例えばXML)を使うことができる。

【0029】本発明の好適実施例において、無線ゲートウェイ14と内容変換アプリケーション16とは、プロキシサーバ上のソフトウェア要素である。図1の各要素の機能は、ハードウェア要素とソフトウェア要素との組み合わせ、あるいはハードウェア要素によって提供しても良い。

【0030】本発明の好適実施例において、図1の各要素の機能は、オブジェクト指向プログラミング技術およびC++プログラミング言語を使うソフトウェアによって提供される。C++の代わりに他のオブジェクト指向プログラミング言語を使っても良い。本発明の好適実施例において、図1の各要素の機能は、非オブジェクト指向プログラミング言語(例えばCプログラミング言語)によって提供しても良い。

【0031】当業者に知られる通り、オブジェクト指向プログラミングは、オブジェクト指向オブジェクトを含むコンピュータソフトウェアを設計するために使われる。これらオブジェクトは、作成が容易であり、変更が経済的であり、再利用可能である。オブジェクト指向オブジェクトは、「オブジェクトデータ」と「オブジェクトサービス」とを含む。オブジェクトサービスは、「オブジェクトメソッド」(「オブジェクトオペレーション」または「オブジェクトファンクション」とも呼ばれる)を介して提供される。オブジェクトメソッドは、一般にオブジェクトが所有する「インスタンスデータ」または「オブジェクトステートデータ」等のプライベートデータ上で動作する。オブジェクトの集まりは、「オブジェクトクラス」と呼ばれる。「オブジェクトタイプ」と呼ばれることもある。オブジェクトクラスは、オブジェクトのセットの行動を記述するためのテンプレートとして働く。オブジェクトの実行は一般に密閉され、公開されず、隠されている。オブジェクトプライベートインスタンスデータは、オブジェクトクラスのオブジェクトメソッドによってのみアクセス可能である。オブジェクトパブリックインスタンスデータは、公開された「オブジェクトインタフェース」を介してアクセスされる。

【0032】本発明の好適実施例に基づく内容変換システム10の各要素の動作環境は、処理システムを含む。この処理システムは、少なくとも1つの高速中央演算処理装置(CPU)とメモリシステムとを含む。以下の本発明の説明は、特に明記しない限り、コンピュータプログラ

ミング分野の実際と、処理システムが実行する動作および命令の働きおよび記号表現とに基づいて行う。このような働きあるいは動作は、「コンピュータ実行」または「CPU実行」と表現する。1つのCPUを用いる例を説明するが、本発明の好適実施例は複数のCPUを用いることもできる。

【0033】前記メモリシステムは、主メモリと二次記憶装置とを含んでも良い。主メモリは、高速ランダムアクセスメモリ(RAM)である。主メモリは、追加のあるいは代替の高速メモリ装置またはメモリ回路を含んでも良い。二次記憶装置は、リードオンリーメモリ(ROM)、光ディスク、磁気ディスク、有機メモリ、揮発性あるいは不揮発性の大容量記憶システム等の長期記憶装置の形態を取ることができる。当業者に明らかな通り、前記メモリシステムは、各種代替要素あるいは代替要素の組み合わせで構成しても良い。

【0034】各要素の働き、および記号で表現する動作や命令は、CPUによる電氣的あるいは生物的信号の処理を含む。電氣的あるいは生物的信号は、データビットを変化させる。すなわち、メモリシステムの記憶場所においてデータビットを保持することは、CPUの動作を再構成あるいは変更することになる。データビットを保持する記憶場所は物理的な場所であり、その記憶場所はそのデータビットに対応する特定の電氣的、磁氣的、光学的、または有機的特性を持つ。

【0035】データビットは、コンピュータあるいはCPUで読取可能な磁気ディスク、光ディスク、有機ディスク、揮発性大容量記憶装置、不揮発性大容量記憶装置等の媒体に格納できる。コンピュータで読取可能な媒体は、共同して動作する媒体、あるいは相互接続した媒体でも良い。これら媒体は処理システムにのみ存在しても良いし、相互接続した複数の遠隔処理システムに分散しても良い。

【0036】本発明の好適実施例の無線ネットワーク装置および他のネットワーク装置は、米国電気電子学会(IEEE)、国際電気通信連合(ITU)電気通信標準化部門、インターネット技術作業部会(IETF)、無線応用プロトコルフォーラム(WAP)等によって提案されている規格に基づく装置であって内容変換システム10と相互作用する装置を含む。IEEE規格は、インターネットアドレス(URL)「www.ieee.org」において参照できる。ITU(以前のCCITT)規格は、「www.itu.ch」において参照できる。IETF規格は、「www.ietf.org」において参照できる。WAPフォーラム規格は、「www.wapforum.org」において参照できる。

【0037】無線装置用内容変換

図2は、内容変換方法24を示すフローチャートである。方法24は、ステップ26において、第1のネットワーク上の第1の無線装置12から、前記第1のネットワーク上の第2のネットワーク装置14にあるオリジナ

表 1

著作権 1998 年スパイグラス社

変換を次のように初期化:

タグ翻訳データテーブル内の各項目について新しいタグ動作辞書およびマークアップタグハッシュテーブルを作成。

タグ動作辞書にタグ翻訳データテーブル項目を挿入。

属性ホワイトリストテーブルの各項目について

前のタグ名が現在の項目のタグ名と同じなら

属性を新しい属性ハッシュセットに挿入

そうでなければ

前のタグ名が空でなければ

新しい属性ハッシュセットをタグ動作辞書に挿入

そうでなければ

新しい属性ハッシュセットを生成

属性を新しい属性ハッシュセットに挿入

前のタグ名を現在の項目のタグ名に設定

最後の新しい属性ハッシュセットをタグ動作辞書に挿入

オリジナル電子文書を次のように前処理:

入力文書を解析し、文書のデータと元のマークアップ言語要素とを代表するトークンのリンクリストを生成。

トークンリストを変更して元のマークアップ言語要素を構文のおよび形式的に正しく表現。

リストの最初から開始。

リストに文書要素が存在する間実行。

トークンが openStartTag なら

トークン名を currentTag にコピー

currentTag が特別の内容タグなら

対応する特別内容フラグを真に設定

currentTag をタグ動作辞書に探索

見つからなければ

currentTag がマークアップタグハッシュテーブルに見つからなければ

現在のトークンを前のトークンに設定

次のトークンから次の closeStartTag トークンまでの全トークンを削除

そうでなければ

次のトークンから次の closeStartTag トークンまでの全トークン(タグ属性)を削除

そうではなくタグに対する動作が removeWithoutContent であれば

現在のトークンを前のトークンに設定

次のトークンから次の closeStartTag トークンまでの全トークンを削除

【 表2 】

表2

テキスト変換	説明
テキスト切り落とし	オリジナル電子文書のサイズは、変換後の形式においても、無線装置への送出に適していない場合がある。これはメモリと帯域幅との制限による。テキストの量を減らすため、内容変換アプリケーション 16 は、特定数の最初の文字または段落と、切り落としたテキストへのリンクとを送出する。この切り落としたテキストへのリンクは、仮想 URL によって行われる。この変換は、所定数の段落要素 (例えば HTML<P>) より大きいと定義された大ブロックのテキストに対して条件付で用いることができる。
サポートされないテキスト要素の除去	第1のマークアップ言語 (例えば HTML) の要素のうち第2の無線マークアップ言語 (例えば WML) がサポートしていないもの (例えば HTML APPLET、BGSOUND、STRIKE 等) は除去する。第1のマークアップ言語における重要でないもの (例えば HTML COMMENTS、META 要素等) は除去し、変換後の文書のサイズを減らす。このような要素は、一般にオリジナル文書の構造またはレイアウトにとって重要でないからである。これは表1の処理中に行う。
サポートされていないテキスト要素タグの除去	第1のマークアップ言語の要素のうち第2の無線マークアップ言語に対応するものがないもの (例えば HTML BODY、ACRONYM、BANNER 等) は除去する。除去した開始および終了要素タグの間のすべての内容は除去しない。これは表1の処理中に行う。
対応テキストマークアップ要素の置換	第1のマークアップ言語の要素および属性のうち、第2の無線マークアップ言語に対応する要素および属性があるものは、適切な要素および/または属性に置換する (例えば HTML の HR は WML の BR に置換する)。第1のマークアップ言語のいくつかの要素および属性は、第2の無線マークアップ言語の複数の要素に置換する。またこの逆もありうる。第1と第2のマークアップ言語の要素のいくつかは同一である (例えばテキストのボールド、イタリック、アンダーラインを表す B、I、U)。これは表1の処理中に行う。

(表2つづき)

	ALIGN="CENTER"/>内容 ALIGN="LEFT"/>へ HTML <TR>内容</TR>から WML 内容 へ HTML <TD>内容</TD>から WML 内容<TAB/>へ HTML <TH>内容</TH>から WML 内容<TAB/>へ 複雑な表はネストされた表を含み、表展開方法を使 って変換する。表展開方法は、表セルを出力書式で 示す。すなわち表セルを括弧で囲み(例えば"["と "]")新しい行を新しい行で始める。
フレーム要素の変換	第1のマークアップ言語のフレーム要素は、フレー ムの全体構造を保存する第2の無線マークアップ言 語に変換する(例えばHTML FRAME ELEMENT は、HTML フレームへのリンクを有するWMLカードと、もし存 在する場合は、HTML NOFRAMES 要素の内容を有する カードとに変換する。FRAME 要素用のHTML SRC 属性 の値は、WML GO 要素に変換する。もし存在すれば、 FRAME 要素用のHTML NAME 属性の値はリンクを特定 するために使用する。そうでなければ、リンクは 「FrameN」と名付ける。ここでNは連続番号であ る。例えばHTML マークアップ<FRAME NAME="name" SRC="URL">はWML マークアップname<GO URL="URL"/>に変換す る。
リンク要素の変換	第1のマークアップ言語におけるリンク要素は、第 2の無線マークアップ言語におけるリンク要素に変 換する。例えばHTML からWML への変換例としてリン クに対して次のような変換を行う。 HTML リンクテキスト から WML リンクテキスト<GO URL="url"/>へ。

ステップ32は、オリジナル電子文書の1つ以上の非テ
キスト要素を第1のマークアップ言語の非テキストフ
ォーマットから第2の無線マークアップ言語の非テキス
トフォーマットに変換する。本発明の好適実施例におい
て、グラフィック画像や画像マップ等の非テキストフ
ォーマットは、第2の無線マークアップ言語で使用可能な
フォーマットに変換される。例えばビットマップ(BMP) 30
画像やJPEG画像は、無線アプリケーションプロトコルで
定義される無線ビットマップ(WBMP)画像に変換される。
BMPおよびJPEG画像フォーマットは、当業者に知られて *

*いる。

【0048】表3は、第1のマークアップ言語の非テキ
スト要素をステップ32において第2の無線マークアッ
プ言語の非テキスト要素に変換するために使われる変換
の例を示す。これらより多くのあるいは少ない非テキ
スト変換を使っても良く、本発明は表3に記載した非テキ
スト変換に限定されない。

【0049】

【表6】

表3

非テキスト変換	説明
画像変換	第1のマークアップ言語で使われるフォーマット の画像を第2の無線マークアップ言語で使用可能 なフォーマットに変換する(例えばJPEG画像を WAPのWBMPフォーマットに変換する)。画像をそ の画像へのリンクに置換すれば、性能を向上させ 表示空間を節約することができる。
クライアント側画像マ ップの変換	第1のマークアップ言語のクライアント側画像マ ップを第2の無線マークアップ言語のマークアッ プに変換し、元の画像にマップされたリンクにユ ーザがアクセスできるようにする(例えばHTML IMG要素またはOBJECT要素のUSERMAP属性は、 HTML文書内のMAP要素を参照する。これに対し てWMLCARD要素が生成される。このWMLCARD要 素は、HTMLHREF属性を持つHTMLMAP内のHTML AREA要素に対するリンクを含み、画像またはオ ブジェクト自体が生成される)。

表3は、非テキスト要素の変換として、画像変換のみを
示すが、本発明は画像変換に限定されず、他の非テキ
スト要素(例えば音声、ビデオ、動画等)にも適用可能であ
る。

表4

```

static tagAttrPair attrWhitelist[] = {
    {"A","HREF"},
    {"A","NAME"},
    {"A","TITLE"},
    {"AREA","ALT"},
    {"AREA","HREF"},
    {"BASE","HREF"},
    {"FIELDSET","TITLE"},
    {"FORM","ACTION"},
    {"FORM","METHOD"},
    {"FRAME","NAME"},
    {"FRAME","SRC"},
    {"IMG","ALT"},
    {"IMG","SRC"},
    {"IMG","USEMAP"},
    {"INPUT","MAXLENGTH"},
    {"INPUT","NAME"},
    {"INPUT","SIZE"},
    {"INPUT","TITLE"},
    {"INPUT","TYPE"},
    {"INPUT","VALUE"},
    {"LINK","HREF"},
    {"MAP","NAME"},
    {"OPTGROUP","TITLE"},
    {"OPTION","TITLE"},
    {"SELECT","NAME"},
    {"SELECT","TITLE"},
    {"SELECT","MULTIPLE"},
    {"TEXTAREA","COLS"},
    {"TEXTAREA","NAME"}
};

// HTML tag translation data table
// table fields:
// tag, tagAction, replacementTag, replStartOnlyFlag, replEndOnlyFlag
static tagTrans tagTransData[] = {
    {"APP", removeWithContent, "", FALSE, FALSE},
    {"APPLET", removeWithContent, "", FALSE, FALSE},
    {"COMMENT", removeWithContent, "", FALSE, FALSE},
    {"DEL", removeWithContent, "", FALSE, FALSE},
    {"PLAINTEXT", removeWithContent, "", FALSE, FALSE},
    {"S", removeWithContent, "", FALSE, FALSE},
    {"SCRIPT", removeWithContent, "", FALSE, FALSE},
    {"STRIKE", removeWithContent, "", FALSE, FALSE},
    {"STYLE", removeWithContent, "", FALSE, FALSE},

    {"ABBR", removeWithoutContent, "", FALSE, FALSE},
    {"ABBREV", removeWithoutContent, "", FALSE, FALSE},
    {"ACRONYM", removeWithoutContent, "", FALSE, FALSE},
    {"AU", removeWithoutContent, "", FALSE, FALSE},
    {"BANNER", removeWithoutContent, "", FALSE, FALSE},
    {"BASEFONT", removeWithoutContent, "", FALSE, FALSE},
    {"BDO", removeWithoutContent, "", FALSE, FALSE},
    {"BGSOUND", removeWithoutContent, "", FALSE, FALSE},
    {"COL", removeWithoutContent, "", FALSE, FALSE},
    {"COLGROUP", removeWithoutContent, "", FALSE, FALSE},
    {"DIR", removeWithoutContent, "", FALSE, FALSE},
    {"DIV", removeWithoutContent, "", FALSE, FALSE},
    {"DL", removeWithoutContent, "", FALSE, FALSE},
    {"FIG", removeWithoutContent, "", FALSE, FALSE},

```

【 表8 】

数「data」は、ルーチンによって変更されるオリジナル画像データを含む。変数「metaData」は、HTTPヘッダデータを含む。表5と同様の機能を使うことにより、他の非テキスト要素(例えば他の画像フォーマット、音声、ビデオ、動画等)を変換できる。本発明は、表5に記載したJPEG画像からWBMP画像への変換ルーチンに限定されない。他の画像フォーマット用のより多くのあるいはより少ない変換機能を使っても良い。

【0054】本発明の他の好適実施例は、JPEG画像を直*

*接WBMP画像に変換する。この時、中間段階を用いない。すなわち、JPEG画像を中間フォーマットに変換し、その中間フォーマットをBMP画像に変換し、そのBMP画像をWBMP画像に変換するという中間段階を省く。かかる実施例においては、単一の変換機能を使う(例えばJPEGtoWBMP(*data))。

【0055】

【表9】

表5

著作権 1998 年スパイグラス社

//JPEG 画像を WML WBMP に変換

//JPEG フォーマット画像を WBMP 画像に変換

//JPEG を内部フォーマットに変換

Convert JPEG_ToInternalFormat(char *data, char *metaData);

//内部フォーマットから BMP フォーマットに変換

//このルーチンは、データを内部フォーマットから BMP フォーマットに変換する

//metaData は次のように変換する

//Content-Length を更新

//Content-Type を image/bmp に変更

ConvertInternalFormatToBMP(char *data, char *metaData)

//BMP フォーマットから WBMP フォーマットに変換

//metaData は次のように変換する

//Content-Length を更新

//Content-Type を image/x-wap.wbmp に変更

ConvertBMP_ToWBMP(char *data, char *metaData)

ステップ34は、変換後のテキスト要素と非テキスト要素とから変換後の電子文書を作成する。本発明の好適実施例において、変換後の電子文書は、WMLおよびWAPプロトコル(例えば画像用のWBMP)を含むWAPによってオリジナルHML文書を変換したものである。オリジナル電子文書は、HML以外のマークアップ言語(例えばcHML、XML、SGML、VRML、VoxML等)を含んでも良く、変換後の電子文書は他の無線マークアップ言語を含んでも良い。

【0056】本発明の好適実施例において、ステップ34は、電子文書を変換するにあたり、無線装置用設定を格納した問い合わせデータベース18を利用する。この実施例において、変換後のテキスト要素および非テキスト要素は、特定の装置設定に合うようにさらに変更される。装置設定の例として、第1の装置タイプは、テキスト表示用として30文字を5行表示するディスプレイを有し、第2の装置タイプは、テキスト表示用として20文字を2行表示するディスプレイを有する。第1の無線装置12の装置タイプは、いずれでも良い。変換後のテ

キスト要素および非テキスト要素は、無線装置のタイプを考慮してステップ34においてさらに変更され、変換された電子文書の最終版を作成する。

【0057】ステップ36において、変換後の電子文書は、オリジナル電子文書に対する要求の応答として、第2のネットワーク装置14から第1の無線装置12に送られる。本発明の好適実施例において、WMLおよび他のWAPプロトコル(例えば画像用のWBMP)を含んだWAP文書が、第2のネットワーク装置14から第1の無線装置12へ、オリジナルHML文書に対する要求の応答として送られる。他のタイプのオリジナル電子文書および他のタイプの変換後の電子文書を用いることも可能であり、本発明はHMLからWAPへの変換に限定されない。

【0058】本発明の好適実施例において、ステップ36は、第1の無線装置が変換後の電子文書のすべてを格納する十分なメモリを持っているかどうかを判断する(例えば要求ヘッダ内の装置タイプから)。もし持っていないならば、変換後の電子文書は複数のサブ文書に分割さ

を複数の文書要素に分割することができる。

【0067】電子文書要素のカテゴリ 分割

この方法は、オリジナル電子文書を作成する時に使われる一般的なカテゴリ (例えばテキスト、リンク、画像、表等)に、オリジナル電子文書を分割する。カテゴリは、リストに表示されて探索される。

【0068】図5は、電子文書のカテゴリ 分割方法62を示すフローチャートである。ステップ64は、第1のマークアップ言語における複数のカテゴリを第1のセットとして選択することにより、オリジナル電子文書を第2の無線マークアップ言語のサブ文書に分割する準備をする。前記第1のセットのカテゴリの各々は、第1のマークアップ言語の1つ以上の文書要素に関連付けられる。ステップ66は、オリジナル電子文書の構造化モデルを作成する (例えば表1において作成されるトークンリストから)。ステップ68は、ループを挿入して、前記第1のセットのカテゴリに基づき、オリジナル電子文書の前記構造化モデルを分解する。ステップ70は、第2の無線マークアップ言語の第1のサブ文書に関連した第1のリストに、1つのカテゴリを追加する。ステップ72は、前記構造化モデル内の1つ以上の文書要素を、前記1つのカテゴリに関連した第2の無線マークアップ言語の次のサブ文書上の次のリストに追加する。ステップ74は、前記第1のサブ文書と前記次のサブ文書との間にリンクを作成する。このリンクは状態情報を含む。ステップ68、70、72、74は、オリジナル電子文 *

表7

次のカテゴリのHTML要素からWMLカードを作成

リンク
テキスト
表
入力フィールド
ページ選択
画像リンク
画像
ナビゲーション

ステップ66は、オリジナル電子文書の構造化モデルを作成する。本発明の好適実施例において、ステップ68は、オリジナル電子文書の文書オブジェクトモデル(DOM)を作成する。このDOMは、ハイパーテキスト電子文書の論理構造を定義し、ハイパーテキスト電子文書のアクセス方法および処理方法を示す。他の構造化モデルを使

表8

変換動作	説明
DOM_CREATE	電子文書から文書オブジェクトモデル(DOM)を作成。
DOM_EVAL	DOMから1つ以上のハイパーテキスト要素を選択するための1つ以上の式を評価。

図6は、HTML用の文書オブジェクトモデル(DOM)76の例の一部を示すブロック図である。HTML用のDOMについての詳細は、ワールドワイドウェブコンソーシアム(W3C)

*書の構造化モデルの最後まで繰り返し、オリジナル電子文書を第2の無線マークアップ言語の複数のサブ文書に分割する。ステップ72を繰り返すことにより、前記次のサブ文書に関連したさらなるサブ文書を作成することもできる。

【0069】本発明の好適実施例において、第1のマークアップ言語はHTMLであり、第2の無線マークアップ言語はWMLである。本発明は、HTMLおよびWMLに限定されない。他のマークアップ言語 (例えばcHTML、SGML、XML、VRML、VoXML等)を使っても良い。HTMLおよびWMLを用いる実施例において、ステップ64は、第1のセットのカテゴリを選択することにより、オリジナルHTML電子文書をWMLカードデッキ内のWMLカードに分割する準備をする。第1のセットのカテゴリの各々は、オリジナルHTML電子文書の1つ以上の文書要素に関連付けられる。(例えばHTML IMG要素は画像カテゴリに関連付けられ、HTML TABLE要素は表カテゴリに関連付けられる)。

【0070】表7は、第1のセットのカテゴリの例としてHTML要素用のカテゴリを示す。本発明は、表7のカテゴリに限定されず、より多くのあるいはより少ないカテゴリを用いることができる。さらに本発明はHTMLカテゴリに限定されず、他のマークアップ言語 (例えばcHTML、SGML、XML、VRML、VoXML等)のカテゴリを使用しても良い。

【0071】

【表11】

※用しても良く、本発明はDOMに限定されない。表8は、ステップ66においてDOMを作成するためのルーチンの例を示す。より多くのあるいはより少ないルーチンを使

【0072】

【表12】

c)文書オブジェクトモデルを参照できる。このw3c DOMの情報は、ワールドワイドウェブのURL「http://www.w3.org」にある。マイクロソフトのダイナミックHTML

表9

HTMLカテゴリ	WML動作
リンク	オリジナルHTML文書からのリンクのリストを提供する。
テキスト	オリジナルHTML文書からHTMLマークアップタグを除いたテキストを提供する。
表	オリジナルHTML文書から表データを提供する。
入力フィールド	HTMLフォームにおけるHTML入力フィールドを示すための入力要素を提供する。
ページ選択	フレームへのリンクの選択リストを提供する。
画像リンク	クライアント側の画像マップによって定義されるリンクの選択リストを提供する。
画像	グラフィック機能を持った無線装置に対して画像へのリンクの選択リストを提供する。
ナビゲーション	状態インデックスを入力として受け取る入力WMLカードを提供する。この入力カードと状態インデックスとによって他のWMLカードへのナビゲーションが提供される。

無線装置に表示される実際の情報量は、その無線装置の解像度(例えば18文字を5行)に依存する。無線装置に送られる情報(例えばWMLカード)の量は、その無線装置のメモリ容量に依存する。状態情報は、複数のWMLカードデッキを接続するために使われる。

【0083】電子文書要素の指定変換

この方法は、オリジナル電子文書から抽出する文書要素を指定するための式を作成する。これら式は、オリジナル電子文書の特定箇所から特定の文書要素を抽出することを可能にする(例えば株価リストから特定の株価を抽出したり、スポーツの点数結果から特定の点数を抽出したりする)。

【0084】図8は、電子文書要素の指定変換方法124を示すフローチャートである。ステップ126は、抽出式のセットを作成し、オリジナル電子文書から1つ以上の文書要素を抽出する準備をする。これら文書要素は、テキスト要素および非テキスト要素を含む。ステップ128は、前記抽出式のセットをオリジナル電子文書に適用して1つ以上の文書要素を抽出する。ステップ130は、抽出した文書要素を用いてオリジナル電子文書を複数の第2の無線マークアップ言語のサブ文書に分割する。

*

表10

最終：1998年4月17日金曜日													
	1	2	3	4	5	6	7	8	9	R	H	E	
Chicago-C	0	0	0	0	0	0	0	2	0	2	5	0	
New York-M	0	1	0	0	0	0	0	0	0	1	3	1	

本発明の好適実施例において、ユーザは無線装置において希望の1つ以上のハイパーテキストを選択できるだけでなく、ユーザに適したフォーマットにその内容を変更できる。例えばユーザがお気に入りのチームが試合に勝ったか負けたかだけを知りたいとする。そしてその無線装置が試合の結果だけを表示するスペースしか持っていないとする。表10は、無線装置に適したボックススコ

【0085】本発明の好適実施例において、内容変換アプリケーション16は、抽出したHTML要素をWML要素に変換する。これは、抽出したHTML要素の特性に基づいて動的にWMLを発生させるか、あるいは抽出したHTML要素をWMLテンプレートに書き込むことによって行う。このWMLテンプレートは、1つ以上のWMLデッキにおける所定セットのWMLカードである。本発明の好適実施例において、ステップ128は抽出式をオリジナル電子文書から作成したDOM76(図6)に適用する。本発明の他の好適実施例において、抽出式は、オリジナル電子文書から作成した文書要素のリストに適用される。

【0086】例えば、野球の好きなユーザがお気に入りのチームの毎日のボックススコアだけを見たいとする。本発明の好適実施例において、ユーザはワールドワイドウェブの電子文書サーバのスポーツページ(例えば新聞のスポーツページ)から希望のボックススコアを選択し、無線装置に毎日表示できる。表10はスポーツページからの野球のボックススコアのフォーマットの例である。

【0087】

【表14】

アのフォーマットを示す。表11は、元のボックススコアの要約であって無線装置の表示に適したものを示す。本発明の好適実施例において、元のボックススコアの要約であって無線装置の表示に適したものがHTMLから抽出され、WMLに変換され、無線装置に転送される。

【0088】

【表15】

表 1 4

タスク	データ収集式の例
箇条書き項目のテキストの獲得	document.all.tags("BL").item(n).children(n).text
表セルからテキストを獲得	document.all.tags("TABLE").item(n).rows(rownum).cells(column).text
1行中の全セルからテキストを獲得。そのテキストを1列に連結。	document.all.tags("TABLE").item(n).rows(rownum).text
1列中の全セルからテキストを獲得。そのテキストを1列に連結。	document.all.tags("TABLE").item(n).rows(rownum).cells(column).text
SELECT 要素からオプションを獲得。テキストは1列に連結。	document.all.tags("FORM").item(n).elements.tags("SELECT").item(n).options.text

例えば、「document.all.tags("TABLE").item(n).children(m).text」というデータ収集式によって表リストの項目nのテキストを獲得するために、オブジェクト指向方法「document」（表12）と方法「all」（表12）と、「tag」（表12）と、HTML表リストタグ「TABLE」と、「item」（表12）と、「children」（表13）と、「テキスト」（表14）とを結合する。データ収集言語の式（例えば表14）は、HTML文書においてネストされているタグを含めて全HTML箇条書きタグからテキストを獲得する。*20

*表14における他のデータ収集式も同様に作成され、HTMLの表およびフォームからテキストを獲得する。

【0093】ハイパーテキスト要素の属性の値を獲得するためのオブジェクト指向方法とデータ収集式の例を表15に示す。これらデータ収集式の例は、表12および13のオブジェクト指向方法を結合している。

【0094】

【表19】

表 1 5

タスク	データ収集式の例
訪問したハイパーリンクの色を獲得	Document.all.tags("BODY").getAttribute("vlink")
アンカーのURLを獲得	Document.all.tags("A").item(n).getAttribute("href")
画像マップ内の領域に関連したURLを獲得	Document.all.tags("MAP").areas(areanum).getAttribute("href")
フォームのURLを獲得	Document.all.tags("FORM").item(n).getAttribute("action")

ハイパーテキスト要素のテキストを検索するためのオブジェクト指向方法とデータ収集式の例を表16に示す。このテキスト検索方法は、指定された文字列を検索し、それを含むすべての要素を返す。その結果、複数の同じテキスト情報が返されることがある。これは、その※

※テキスト検索方法がDOM7 6 (図6)における要素階層の底で使用されない場合である。

【0095】

【表20】

表 1 6

タスク	データ収集式の例
フレーム内において文字「Contents」を含む要素を検索。	Document.all.tags("FRAME").grep("CONTENTS")

このデータ収集式の例は、無線装置の表示用に情報を抽出するために使うことができる。抽出式はデータベース18 (図1)に格納することができる。図9は、図8の電子文書要素の指定変換方法124を示すブロック図132である。オリジナルHTML電子文書134は、スポーツの点数（例えば表10）を含む表136を含む。抽出式138を表136に適用し、WMLデッキ142のWMLカード140を作成する。これは元の表136のデータのサブセット（例えば表11）を含む。方法124 (図8)を方法24 (図2)のステップ28で使うことにより、オリジナル電子文書を複数の文書要素に分割できる。

40 【0096】本発明の好適実施例の使用例

図10および11は、本発明の好適実施例の使用例を示す。ただし図10および11は例示が目的であって、本発明はこの例示に限定されない。図10は、スクリーン表示142であり、ワールドワイドウェブからのHTML電子文書の出力例を示す。この電子文書の例は、スクリーン表示142の「最新ニュース」(LATEST NEWS)の欄に目的のテキスト144がある。表17は、図10のスクリーン表示142を表示するために使われるHTMLの例の一部である。表17に示したHTMLの例において、「最新ニュース」(LATEST NEWS)欄のテキスト144は、HTML

50

(表17のつづき)

```

<!-- cell 1 -->
<a href="/newsflash/releases/113098swwin98.html" class=SIDEBAR>Alga To Partner
with Spyglass on Advanced TV Font Rendering <IMG SRC="/images/Arrow_r2.gif" WIDTH=9
HEIGHT=9 BORDER=0 hspace=3 align=middle ALT="Go"></a><br><br>
<a href="/newsflash/releases/111998swpccomputing.html" class=SIDEBAR>Spyglass
Ships New Version of Device Mosaic<IMG SRC="/images/Arrow_r2.gif" WIDTH=9 HEIGHT=9
BORDER=0 hspace=3 align=middle ALT="Go"></a><br><br>
<a href="/newsflash/releases/111698wince.html" class=SIDEBAR>Device Mail 3.1
Embedded Email Client Announced<IMG SRC="/images/Arrow_r2.gif" WIDTH=9 HEIGHT=9
BORDER=0 hspace=3 align=middle ALT="Go"></a><br><br>
</font>
</DIV>
<!-- end Latest News items -->
</td>
<td width=35 valign=top>
<br>
</td>
</tr>
<tr>
<td width=200 valign=top colspan=4>
<br>
<br>
</td>
</tr>
</table>
...
</html>

```

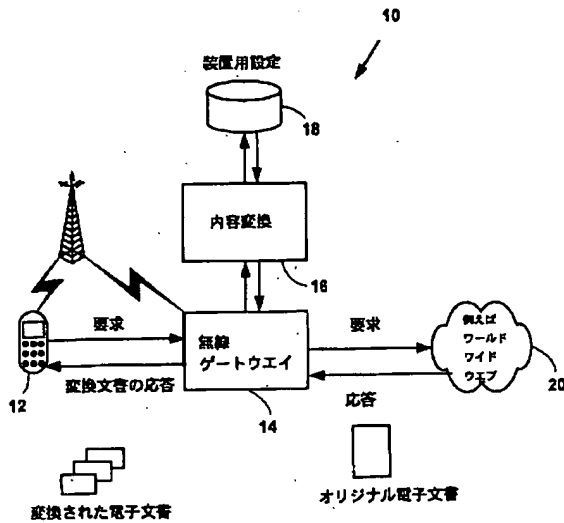
図11は、図10のスクリーン表示から変換した出力を、第1の無線装置12に表示した例を示すブロック図146である。スクリーン表示142(図10)の「最新ニュース」(LATEST NEWS)欄のテキスト144は、第1の無線装置12(図11)の無線ディスプレイ148に表示されている。図10の電子文書表示は、表17のHTML文書に基づいている。無線ディスプレイ148の変換出力表示は、WML文書のWMLカードに基づいている。図10のスクリーン表示142のHTML文書は、本発明の好適実施例に基づく1つ以上の方法およびシステムによって変換され、WMLカードとして第1の無線装置12の無線デ

ィスプレイ148に表示される。スクリーン表示142の「最新ニュース」(LATEST NEWS)欄の他のテキストを追加のWMLカードによって表示することもできる。【0098】スクリーン表示142(図10)の「最新ニュース」(LATEST NEWS)欄のテキスト144を見る場合、テキスト144(図10)は表の一部であり、TABLE要素はオリジナル電子文書の第1のセットの文書要素である。方法24(図2)のステップ28において、方法24と方法40(図3)とを併用し、図11の無線装置ディスプレイ148の出力を作成できる。この実施例において、図11の無線装置ディスプレイ148の出力は、第

43

- 10 内容変換システム
12 無線装置
14: 無線ゲートウェイ

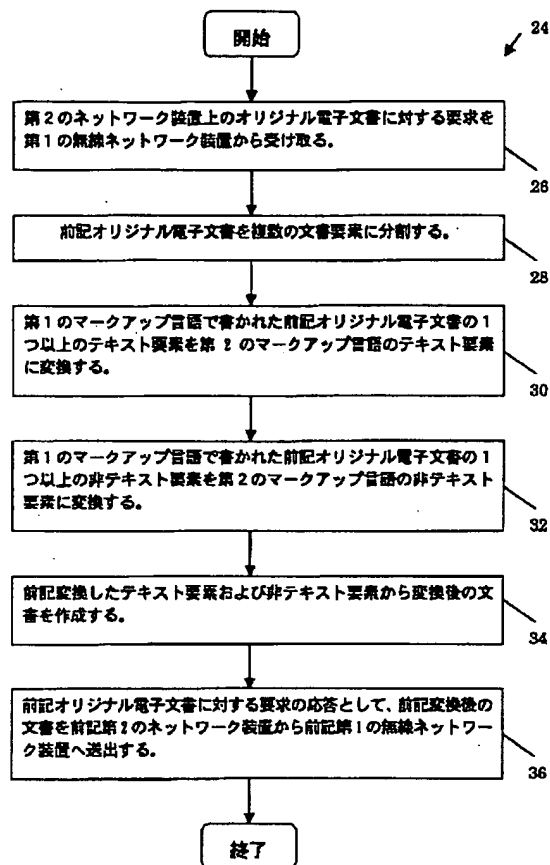
【図1】



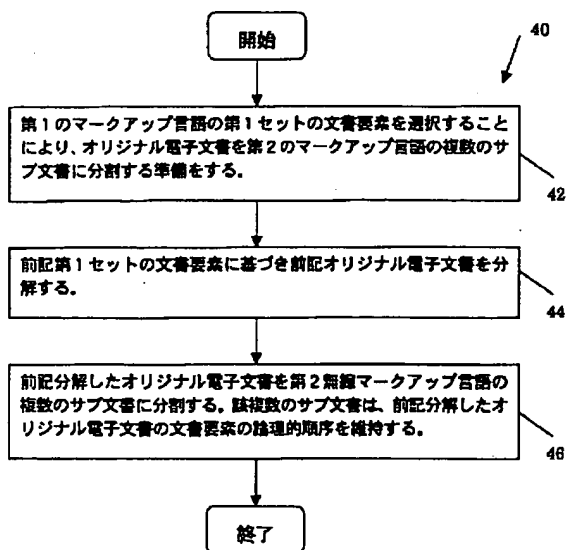
44

- 16: 内容変換アプリケーション
18: 装置用設定データベース
20 ワールドワイドウェブ (WWW)

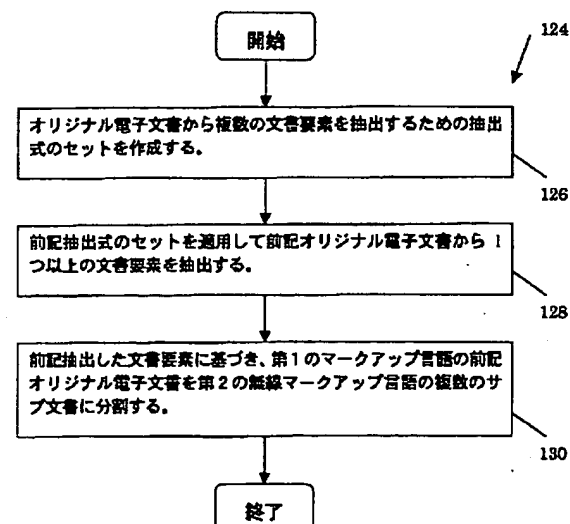
【図2】



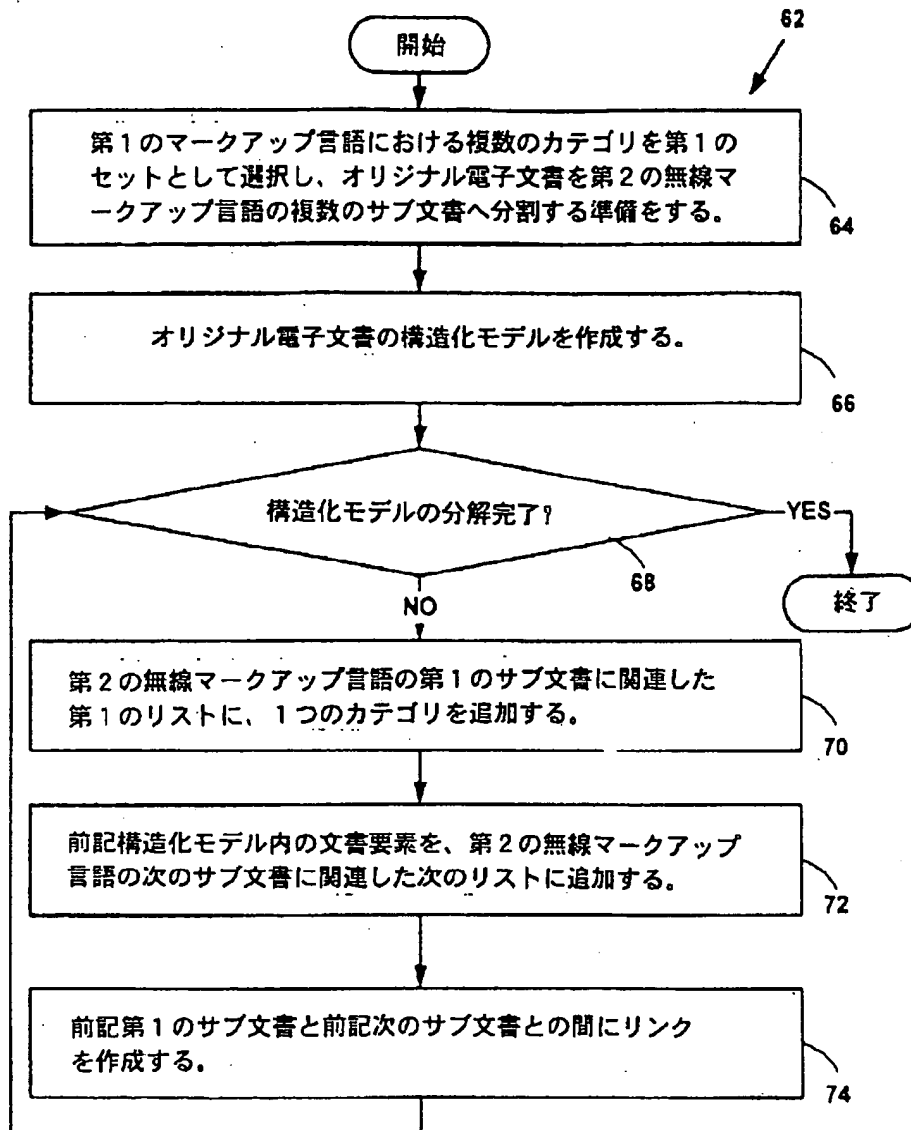
【図3】



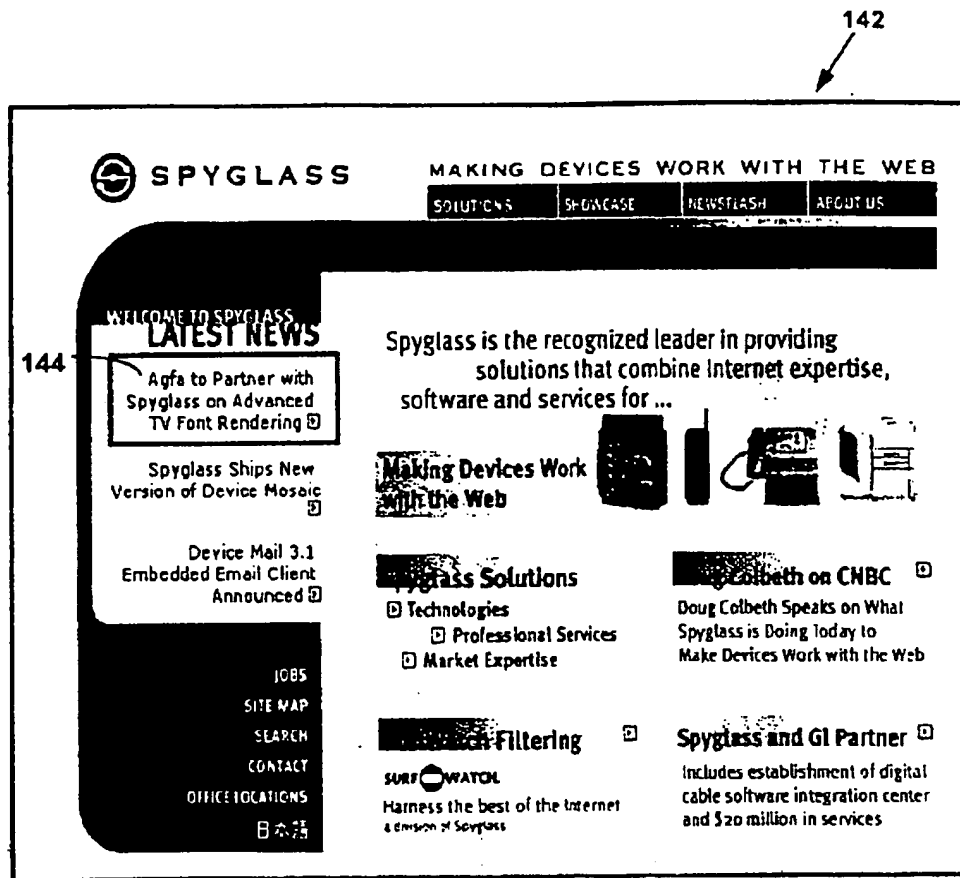
【図8】



【 図5 】



【 図10 】



フロント ページの続き

(72)発明者 ケネス ビー. ホール
 アメリカ合衆国 60525 イリノイ州 ラ
 グランジュ サウス テンス アヴェニュー
 ー 525